Joint stochastic inversion of geophysical data for reservoir parameter estimation

Jinsong Chen and G. Michael Hoversten, Lawrence Berkeley National Laboratory*

Summary

A stochastic model is developed to estimate porosity (ϕ) and water saturation (S_w) using multiple sources of information, including borehole ϕ and S_w measurements, seismic P- and S-wave travel time, and inverted electrical conductivity (σ). Within the stochastic framework, both reservoir parameters and geophysical attributes at unsampled locations in the interwell volumes are considered as random variables and are estimated simultaneously by conditioning to available data. The focus of the inversion process is not on finding a best-fitting solution of the unknown parameters as with conventional methods, but rather on sampling from the joint probability density function of the unknowns. From the samples, various statistics of each variable can be inferred, such as mean, variance, confidence intervals, and even probability density functions. A synthetic case study based on well log data is presented. Results show that the stochastic inverse provides not only overall better estimates of ϕ and S_w but also information about uncertainty in the estimation that cannot be obtained using conventional methods.

Introduction

Conventional methods for reservoir parameter estimation using multiple sources of geophysical data include two major steps. First, each type of geophysical measurements is inverted in isolation to produce an image of geophysical properties. Then the various properties are combined to estimate reservoir parameters based on petrophysical or rock physical models obtained from borehole data. This approach has limitations due to the inability to quantify uncertainty in the estimated parameters and due to the lack of information sharing among various types of data in the inversion process.

Stochastic inverse methods have been used recently to combine multiple sources of geophysical data. For example, Bosch and McGaughey (2001) developed a statistical model to jointly invert gravity and magnetic data for lithology estimation with geologic constraints. Eidsvik et al. (2002) developed a Bayesian method to estimate facies, porosity, fluid saturation, and density, using seismic AVO attributes and various types of borehole measurements. In those studies, reservoir parameters as well as geophysical attributes at unsampled locations were considered as random variables, and the Markov chain Monte Carlo (MCMC) methods (Gilks et al., 1996) were used to draw samples from their joint probability density functions (pdfs). The petrophysical or rock physical relationships between the reservoir parameters and the geophysical attributes were enforced. Those methods provide a means to characterize uncertainty in parameter estimation and allow multiple sources of information to be shared in the inversion process.

We develop a stochastic model to estimate ϕ and S_w in this study, using borehole ϕ and S_w measurements, seismic P- and S-wave travel time, and inverted σ (Newman, 1995) from crosswell EM data. Figure 1 shows the relationships between the reservoir parameters and the geophysical attributes. Unlike conventional inversion, our stochastic inversion of seismic P-wave velocity links P-wave travel time, S-wave travel time, and the inverted σ. Estimation of ϕ and S_w depends on borehole measurements, P- and S-wave travel time, and inverted σ. Although σ is considered as data in the current model, our ultimate goal is to consider it as a random variable, conditioned to crosshole EM amplitude and phase measurements.

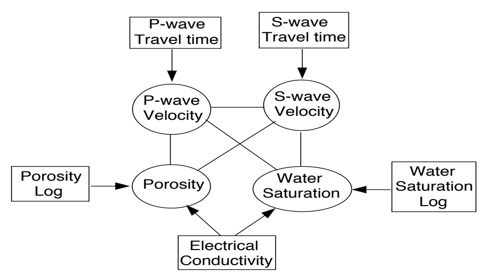

Figure 1: Relationships between reservoir parameters and geophysical attributes. The rectangles represent data, the circles represent random variables, the arrows represent dependence relationships, and the lines represent mutual dependence relationships.

Stochastic Model

<u>Bayesian framework</u>

We combine various types of data using a Bayesian framework. Consider a two-dimensional cross section, which is divided into m pixels and passed through by n seismic rays. Let vectors $\mathbf{\Phi}$, $\mathbf{S_w}$, $\mathbf{S_p}$, and $\mathbf{S_s}$ be the unknown ϕ, S_w, and P- and S-wave slowness at the m pixels. Let vectors $\mathbf{t_p}$ and $\mathbf{t_s}$ be the P- and S-wave travel time of the n rays, and vector $\boldsymbol{\sigma}$ be the inverted σ at the m pixels. Let τ_p, τ_s, and τ_c be the inverse variances of the measurement errors of the P- and

S-wave travel time and the inverted electrical conductivity. Within the stochastic framework, all the unknowns are considered as random variables. Instead of seeking the best-fitting values of the variables as with a deterministic inverse, our goal is to fully characterize uncertainty of each variable, such as its mean, variance, and probability density function.

We begin by deriving the joint probability density function of all the unknown variables. According to the Bayes' theorem (Stone, 1995), the joint conditional pdf (also referred to as the posterior pdf) given P- and S-wave travel time and inverted electrical conductivity can be written as:

$$\begin{aligned} f(\mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s},\tau_p,\tau_s,\tau_c \mid \mathbf{t_p},\mathbf{t_s},\boldsymbol{\sigma}) \propto \\ f(\mathbf{t_p},\mathbf{t_s},\boldsymbol{\sigma} \mid \mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s},\tau_p,\tau_s,\tau_c)\cdot \\ f(\mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s},\tau_p,\tau_s,\tau_c). \end{aligned} \quad (1)$$

The first term on the right side is referred to as the likelihood function, and the second term is referred to as the prior pdf.

Conventional inversion methods minimize the misfit between modeled and measured data using the least square estimation method, which is equivalent to the maximum likelihood method when the measurement errors have a multivariate Gaussian distribution. The prior information is either ignored or partially combined implicitly. In this study, we explore the posterior pdf, which explicitly combines information from both the likelihood function and the prior. We draw samples from the pdf using the Markov chain Monte Carlo methods and fully characterize uncertainty of each variable by evaluating those samples.

Likelihood function

The likelihood function is the link between unknown random variables and given geophysical data. Consider that P- and S-wave travel time has a direct connection with their corresponding slowness and measurement errors, and σ is a function of ϕ, S_w, and its measurement error, the likelihood function in Equation 1 can be written as follows:

$$\begin{aligned} f(\mathbf{t_p},\mathbf{t_s},\boldsymbol{\sigma} \mid \mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s},\tau_p,\tau_s,\tau_c) \propto \\ f(\mathbf{t_p} \mid \mathbf{S_p},\tau_p) f(\mathbf{t_s} \mid \mathbf{S_s},\tau_s) f(\boldsymbol{\sigma} \mid \mathbf{\Phi},\mathbf{S_w},\tau_c). \end{aligned} \quad (2)$$

Each term on the right side of the above equation can be obtained using their corresponding forward models or petrophysical relations. Let M be the forward model of seismic tomographic data. The measured P- and S-wave travel time thus can be written as:

$$\mathbf{t_p} = M(\mathbf{S_p}) + \varepsilon_p, \quad (3)$$

$$\mathbf{t}_s = M(\mathbf{S}_s) + \varepsilon_s. \quad (4)$$

Vectors ε_p and ε_s are the measurement errors of P- and S-wave travel time, and both have the multivariate Gaussian distribution with zero mean. Consequently, $f(\mathbf{t_p} \mid \mathbf{S_p},\tau_p)$ is the pdf of the multivariate Gaussian distribution with mean $M(\mathbf{S_p})$ and covariance matrix $\tau_p^{-1/2}\mathbf{I_n}$, and $f(\mathbf{t}_s \mid \mathbf{S}_s,\tau_s)$ is the pdf of the multivariate Gaussian distribution with mean $M(\mathbf{S}_s)$ and covariance matrix $\tau_s^{-1/2}\mathbf{I_n}$, where $\mathbf{I_n}$ is the n dimensional identity matrix.

The likelihood function of $\boldsymbol{\sigma}$ is obtained using Archie's law (Archie, 1942). In log space, the relationship between porosity, water saturation, and electrical conductivity is linear and given by:

$$\log(\boldsymbol{\sigma}) = w_0 + w_1 \log(\mathbf{\Phi}) + w_2 \log(\mathbf{S_w}) + \varepsilon_c. \quad (5)$$

Coefficients w_0, w_1, and w_2 are determined from well log data. Vector ε_c is the measurement or inversion error of log electrical conductivity and has the multivariate Gaussian distribution with zero mean and covariance matrix $\tau_c^{-1/2}\mathbf{I_m}$, where $\mathbf{I_m}$ is the m dimensional identity matrix.

Prior pdf

The prior pdf in Equation 1 summarizes the information that is not included in the above likelihood functions, such as borehole ϕ and S_w measurements. As measurement errors of P- and S-wave travel time and electrical conductivity are often independent of $\mathbf{\Phi}$, $\mathbf{S_w}$, $\mathbf{S_p}$, and $\mathbf{S_s}$, we can write the prior pdf as follows:

$$f(\mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s},\tau_p,\tau_s,\tau_c) = f(\mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s}) f(\tau_p,\tau_s,\tau_c). \quad (6)$$

The first term on the right side of Equation 6 is derived from borehole data. For example, we can expand it as follows:

$$\begin{aligned} f(\mathbf{\Phi},\mathbf{S_w},\mathbf{S_p},\mathbf{S_s}) = \\ f(\mathbf{S_s} \mid \mathbf{S_p},\mathbf{S_w},\mathbf{\Phi}) f(\mathbf{S_p} \mid \mathbf{S_w},\mathbf{\Phi}) f(\mathbf{S_w} \mid \mathbf{\Phi}) f(\mathbf{\Phi}). \end{aligned} \quad (7)$$

Each individual conditional pdf is obtained from borehole data using regression methods. The prior pdf of $\mathbf{\Phi}$ is obtained using geostatistical kriging by conditioning to borehole porosity measurements. For the second term on the right side of the equation, we often have less information. In practice, we usually choose a prior for each of the variables that has minimal effect on the analysis (Gilks et al., 1996).

Sampling Method

The key to success of this stochastic inverse lies in the use of the Markov chain Monte Carlo (MCMC) methods. These methods have recently been emerged as the main tools for solving problems involving a large number of random variables. As the posterior pdf in Equation 1 is very complicated, analytical methods cannot be used. Instead, we use a MCMC method to draw many samples for each variable. The method is different from the traditional Monte Carlo methods, which draw independent samples from conditional distributions. The MCMC method draws dependent samples by running a cleverly constructed Markov chain from given starting values (Gilks et al., 1996). The initially drawn samples depend on the starting values. Therefore, they are not used for inference in order to avoid the bias in choosing the starting values. The remaining samples are independent of the starting values and thus all are used for inference. The total number of runs needed and the number of the initial values discarded are determined by convergence diagnosis methods (Gelman and Rubin, 1992).

We use the Gibbs sampling method (Gilks et al., 1996) to generate Markov chains in this study. The sampling method includes the following steps: (1) Defining likelihood functions for given data and prior pdfs for random variables, (2) Deriving conditional pdfs for random variables, and (3) Drawing samples from the conditional pdfs. The above sampling process is repeated so that many samples are obtained for each random variable.

Case Study

We construct a synthetic dataset to demonstrate the performance of the stochastic inverse. The dataset was generated using well log data collected from two wellbores in the Lost Hills oil field in Southern California by Chevron Petroleum Company (Hoversten et al., 2002). A synthetic reservoir interval was constructed between the two wells, separated by 150m. The ϕ and S_w of the reservoir are shown in Figure 2. Our goal is to estimate ϕ and S_w within the reservoir using direct measurements of ϕ and S_w at each well, seismic P- and S-wave travel time, and inverted σ on the interwell cross section.

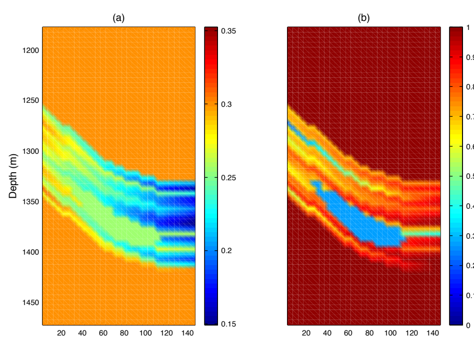

Figure 2: True ϕ (a) and S_w (b)

We consider a cross section, which passes through the two wellbores at depths 1175m to 1475m. We divide the two-dimensional domain (150m × 300m) into 1800 pixels with the dimensions of 5m × 5m. In addition to borehole ϕ and S_w data, we also collect seismic P- and S-wave travel time from 1600 source-receiver combinations between the two wells, and electrical conductivity from inversion of a simulated crosswell electromagnetic experiment. Gaussian random noises were added to each type of data. The EM data are assumed to have 3% noises. The P-wave travel time error is equal to the 3% of the averaged P-wave travel time, and the S-wave travel time error is equal to the 5% of the averaged S-wave travel time.

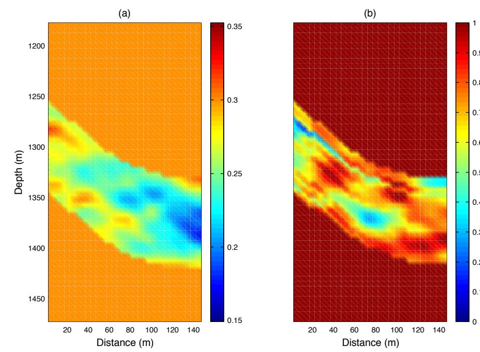

Figure 3: Estimated ϕ (a) and S_w (b) using a regression model based on the inverted P-and S-wave velocity and the inverted σ.

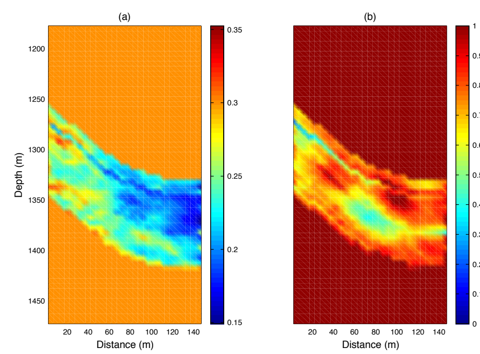

Figure 4: Estimated ϕ (a) and S_w (b) using the stochastic inverse based on the P- and S-wave travel time and the inverted σ.

Figure 3 shows the estimated values of ϕ and S_w using the conventional approach. In this approach, P- and S-wave travel time was inverted separately to get P- and S-wave slowness using the ART algorithm (Peterson, 1985). The models for predicting ϕ and S_w from P- and S-wave slowness and σ were obtained from data at the two wellbores using a multivariate regression (Stone, 1995). For ϕ, the estimated results capture the main patterns of the true values, that is, the ϕ reduces from the left side to the right side, but they are much smoother than the corresponding true values. For S_w, the estimated results correctly predict the existence of a low S_w zone in the lower middle part of the reservoir, but underestimate the values at the left and the upper right sides of the reservoir.

Figure 4 shows the estimated ϕ and S_w using the stochastic inverse based on borehole ϕ and S_w measurements, P- and S-wave travel time, and inverted σ. Compared to Figure 2 and Figure 3, the estimated values of both ϕ and S_w are better than the corresponding estimated values using the conventional approach in terms of overall patterns and resolution. A quantitative comparison between the estimated ϕ and S_w using the traditional and stochastic methods is summarized in Table 1. The stochastic inverse gives smaller ranges of the estimated values, smaller prediction errors, and smaller mean differences over all the pixels.

Table 1. Comparison of estimated ϕ and S_w with the corresponding true values. The true ϕ is in the range from 0.15 to 0.31, and the true S_w is in the range from 0.22 to 1.0.

	Traditional Method	Stochastic Method
Porosity (ϕ)		
Range of values	(0.17,0.32)	(0.15,0.32)
Range of errors	(-0.07,0.14)	(-0.09,0.05)
Mean difference	0.021	0.017
Water Saturation (S_w)		
Range of values	(0.16,1.45)	(0.22,1.05)
Range of errors	(-0.54,0.63)	(-0.39,0.54)
Mean difference	0.155	0.117

Conclusions

We developed a stochastic model for ϕ and S_w estimation using multiple sources of geophysical data. The model has the following advantages over the deterministic method:

- Provide a means to quantify uncertainty in reservoir parameter estimation. Unlike the deterministic approach, which gives only one value for each unknown parameter, the stochastic inverse draws many samples from the joint conditional pdf of each variable. From the samples, we can calculate the mean, variance, confidence interval, and even probability density function of the variable.
- Allow various types of measurement and model errors to be considered in parameter estimation. The deterministic method typically cannot account for model errors and underestimates the effects of measurement errors. Our method considers both measurement and model errors as random variables, and they are estimated in the same way as the reservoir parameters.
- Provide a better way to integrate multiple sources of information. The deterministic approach integrates various types of geophysical data after each individual inversion. This excludes information sharing among those data in the inversion process. The stochastic inverse estimates unknown reservoir parameters and geophysical attributes simultaneously.

Acknowledgements

The authors thank John Peterson from the Lawrence Berkeley Laboratory for assistance in deterministic tomographic inversion and the Chevron Petroleum Company for providing data for this study. This work was funded by the Assistant Secretary for Fossil Energy, National Petroleum Office of the U.S. Department of Energy, under contract DE-AC03-76SF0098 as part of the National Gas and Oil Technology Partnership program.

References

Archie, G. E., 1942, The electrical resistivity log as an aid in determining some reservoir characteristics: Trans. Am. Inst. Mech. Eng., **146**, 54-62.

Bosch, M., 1999, Lithologic tomography: From plural geophysical data to lithology estimation: J. Geophy. Res., **104**, 749-766.

Eidsvik, J., Henning, O., Mukerji, T., Mavko, G., and Avseth, P., 2002, Seismic reservoir prediction using Bayesian integration of rock physics and Markov random fields: A North Sea example: The Leading Edge, 290-294.

Gelman, A. and Rubin, D., 1992, Inference from iterative simulation using multiple sequences: Statistical Science, **7**, 457-511.

Gilks, W., Richardson, S., and Spiegelhalter, D., 1996, Markov Chain Monte Carlo in Practice: Chapman & Hall/CRC, New York.

Hoversten, M., Gritto, R., Washbourne, J., and T. Daily, 2002, Fluid saturation and pressure prediction in a multi-component reservoir by combined seismic and electromagnetic imaging: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys.

Newman, G. A., 1995, Crosswell electromagnetic inversion using integral and differential equations: Geophysics, **60**, 899-911.

Peterson, J. E., Paulsson, B. N., and McEvilly, T. V., 1985, Applications of algebraic reconstruction techniques to crosshole seismic data: Geophysics, **50**, 1566-1580.

Stone, C. J., 1995, A Course in probability and statistics: Duxbury, New York.